Journal of Heredity 2008:99(2):112–124
doi:10.1093/jhered/esm104
Advance Access publication January 24, 2008

Utility of EST-Derived SSRs as Population Genetics Markers in a Beetle

KYUNG SEOK KIM, SUSAN T. RATCLIFFE, B. WADE FRENCH, LEI LIU, AND THOMAS W. SAPPINGTON

From The USDA-ARS, Corn Insects and Crop Genetics Research Unit, Genetics Laboratory, Iowa State University, Ames, IA 50011; (Kim and Sappington); the University of Illinois, Department of Crop Science, S-316 Turner Hall, 1102 S. Goodwin Avenue, Urbana, IL 61801 (Ratcliffe); the USDA-ARS, North Central Agricultural Research Laboratory, 2923 Medary Avenue, Brookings, SD, 57006 (French); the University of Illinois, Biotechnology Center, 330 ERML, 1201 Gregory Drive, Urbana, IL 61801 (Liu).

Address correspondence to T. W. Sappington at the address above, or e-mail: tom.sappington@ars.usda.gov.

Abstract

Microsatellite, or simple sequence repeat (SSR), loci can be identified by mining expressed sequence tag (EST) databases, and where these are available, marker development time and expense can be decreased considerably over conventional strategies of probing the entire genome. However, it is unclear whether they provide information on population structure similar to that generated by anonymous genomic SSRs. We performed comparative population genetic analyses between EST-derived SSRs (EST-SSRs) and anonymous SSRs developed from genomic DNA for the same set of populations of the insect *Diabrotica virgifera*, a beetle in the family Chrysomelidae. Compared with noncoding, nontranscribed regions, EST-SSRs were generally less polymorphic but had reduced occurrence of null alleles and greater cross-species amplification. Neutrality tests suggested the loci were not under positive selection. Across all populations and all loci, the genomic and EST-SSRs performed similarly in estimating genetic diversity, F_{IS}, F_{ST}, population assignment and exclusion tests, and detection of distinct populations. These findings, therefore, indicate that the EST-SSRs examined can be used with confidence in future genetic studies of *Diabrotica* populations and suggest that EST libraries can be added as a valuable source of markers for population genetics studies in insects and other animals.

Because they are common, hypervariable, and have co-dominant inheritance, microsatellites, or simple sequence repeats (SSRs), are widely used by population geneticists for testing ecological and evolutionary hypotheses in natural populations (Zhang and Hewitt 2003; Selkoe and Toonen 2006; Ellis and Burke 2007). However, SSRs are, as a rule, species-specific markers which must be developed de novo for each species, mainly because they commonly occur in noncoding regions of the genome which are not highly conserved (Zane et al. 2002). Significant progress has been made in developing more efficient methods of obtaining new SSRs (Zane et al. 2002), but these markers remain expensive and time-consuming to isolate (Squirrel et al. 2003; Ellis and Burke 2007).

Over the past decade, genomics technology has provided researchers with access to vast amounts of sequence information through public and private databases (Varshney et al. 2005; Bouck and Vision 2007; Ellis and Burke 2007; Kong et al. 2007). As an alternative to the conventional strategy of probing repeat-enriched genomic libraries for anonymous SSRs, large numbers of novel SSRs can be isolated with comparatively minor effort simply by in silico mining of expressed sequence tag (EST) databases (Liu, Karsi, Dunham 1999; Liu, et al. 1999; Coulibaly et al. 2005; Perez et al. 2005; Varshney et al. 2005; Chen et al. 2006; von Stackelberg et al. 2006; Yi et al. 2006; Bouck and Vision 2007; Kong et al. 2007). Developing large numbers of ESTs from a species has become routine, and there are many characteristics of EST-derived SSRs that make them valuable as genetic markers. These include presence in large numbers, high levels of polymorphism compared with many other types of genetic markers, locus-specific codominant inheritance, repeatability and clarity of scoring, and enhanced transferability across related species.

Indeed, EST-SSRs have been widely used to construct high-density linkage maps (Fraser et al. 2004; Prasad et al. 2005; Chen et al. 2006; Marcel et al. 2007; Slate et al. 2007; Stein et al. 2007) and have proved useful for detecting the signature of divergent selection or exploring adaptive behavior in a range of species (Li et al. 2002; Vigouroux et al. 2002; Luikart et al. 2003; Vasemagi et al. 2005; Yatabe et al. 2007). These together with numerous other studies (Cho et al. 2000; Cordeiro et al. 2001; Eujayl et al. 2001, 2004; Leigh et al. 2003; Liewlaksaneeyanawin et al. 2004;

Coulibaly et al. 2005; Woodhead et al. 2005; Grillo et al. 2006; Pashley et al. 2006) indicate that EST-SSRs generally have fewer null alleles, greater cross-species amplification, and less allelic variability than genomic SSRs. These properties are related to the higher level of conservation of DNA sequences in coding than in noncoding regions of the genome (Wren et al. 2000; Li et al. 2002; Woodhead et al. 2005; Bouck and Vision 2007; Slate et al. 2007; Yatabe et al. 2007).

The availability of EST-SSRs for population genetics studies would be an important option for a number of taxa, where development of conventional SSRs has proved problematic. For example, null alleles are more of a problem for some animal taxa than others (Liewlaksaneeyanawin et al. 2002). A low frequency of genomic SSRs, a history of duplication/multiplication events within the genome, and frequent crossing over between nonhomologous SSRs leading to the exchange of flanking regions have made developing usable SSRs very difficult for some animal taxa, including Lepidoptera (Neve and Meglecz 2000; Meglecz et al. 2004; Zhang 2004), *Aedes* mosquitoes (Fagerberg et al. 2001), mites (Navajas et al. 1998), ticks (Fagerberg et al. 2001), nematodes (Grillo et al. 2006, Johnson et al. 2006), bivalve molluscs (Cruz et al. 2005), and birds (Primmer et al. 1997; Neff and Gross 2001). Efficiency of isolation is the main factor limiting use of SSRs in these species, and new methods are needed (Zhang 2004). Because the stability of the DNA sequences flanking SSRs from transcribed regions reduces both null alleles and duplication events, EST database mining may provide a better means for obtaining useful SSRs for problematic taxa, as was demonstrated recently for a parasitic nematode (Grillo et al. 2006).

Employment of EST-SSRs in population genetics applications has been limited to a handful of studies in plants (Leigh et al. 2003; Woodhead et al. 2005; Ellis et al. 2006; Yatabe et al. 2007) and animals (Walton et al. 2004; Coulibaly et al. 2005; Miller et al. 2005; Perez et al. 2005; Grillo et al. 2007). The hesitation to use EST-SSRs in population studies is mainly because some loci may be under selective pressure or subject to background selection caused by genetic hitchhiking and population genetics applications require selectively neutral markers (Ford 2002; Li et al. 2002; Selkoe and Toonen 2006; Ellis and Burke 2007). On the other hand, allozymes have demonstrated their robustness as population genetics markers in myriad studies since the 1960s (Avise 2006; Behura 2006), even though they suffer from the same potential problems arising from exposure to selection. Ellis and Burke (2007) point out that although selection can be expected to be operating on a small percentage of EST-SSRs, this drawback can be largely overcome by use of a sufficient number of markers, which minimizes the influence of a nonneutral locus on estimates of population parameters. Furthermore, there are a number of statistical methods available for flagging loci under apparent selection (Ewens 1972; Watterson 1978; Slatkin 1994; Beaumont and Nichols 1996), which can then be removed from the analysis. Thus, there is reason to think that EST-SSRs can be useful in characterizing the genetic structure of populations.

So far, direct comparisons of the performance of EST-SSRs and genomic SSRs in inferring population genetic structure have been reported in only 2 studies, both of plants. In a study of the alpine lady-fern, Woodhead et al. (2005) concluded that many polymorphic EST-SSRs behave as effectively neutral markers and can provide an informative signature for plant population genetic structure. Yatabe et al. (2007) reported that the estimate of F_{ST} between 2 hybridizing species of sunflower was higher when based on EST-SSRs than when based on anonymous SSRs, attributable to the lower mean allelic diversity of the EST-SSRs. In addition, a few EST-SSRs appeared to be under selection. Nevertheless, both types of markers led to comparable conclusions about population structuring and gene flow. To our knowledge, there have been no direct comparisons in animals of EST and genomic SSRs for estimating genetic variation, population genetic structuring, and population assignment. Among the few reports of using EST-SSRs to study animal populations, only Coulibaly et al. (2005) noted a manifest shortcoming. The authors concluded that EST-SSRs appeared to be inferior to genomic SSRs in addressing phylogenetic relationships among salmonid fish species because of their higher level of conservation. Nevertheless, the authors considered the EST-SSR markers a valuable resource for population and evolutionary genetic studies because of their high level of interspecific transferability.

In this study, we screened large numbers of SSRs from an EST database of western corn rootworm (*Diabrotica virgifera virgifera*), a chrysomelid beetle, and a major insect pest of corn in North America and Europe (Sappington et al. 2006). Based on the same criteria for selecting genomic SSRs, 38 EST-SSRs were selected for evaluation of their transferability across corn rootworm species, the occurrence of null alleles, and the level of genetic variability. More importantly, we directly compared the performance of EST-SSRs with previously developed genomic SSRs (Kim and Sappington 2005b) in characterizing population genetic structure and population assignment for the same set of individuals sampled from 6 populations.

Materials and Methods

Sample Collection

Western corn rootworm adults were sampled from 5 widely separated geographic locations across 5 states of the United States: Ankeny, IA, in 2003 (IA); Champaign, IL, in 2003 (IL); Dodge City, KS, in 2003 (DKS); Cobleskill, NY, in 2003 (NY); and New Deal, TX, in 2004 (TX). Details of western corn rootworm sample collection are described in a previous report (Kim and Sappington 2005a). Mexican corn rootworms (MCRs), *Diabrotica virgifera zeae*, were collected near Goliad, Texas in 2004. Western and Mexican corn rootworms differ morphologically and are considered subspecies. These subspecies are partially reproductively isolated because cytoplasmic incompatibility generated by the rickettsial bacterium *Wolbachia* (Giordano et al. 1997) permits productive mating only between *D. v. virgifera*

Table 1. Characteristics of 17 SSRs derived from *Diabrotica virgifera virgifera* ESTs in a survey of 60 individuals sampled from a central Iowa population

Locus	Primer sequences (5′-3′)	Repeat motif	Number of alleles (evidence of null allele)[a]	H_O	H_E	P^b	GenBank accession number
DVV-ED1	F: ACCACCACAGAGCTGGAATC R: GTCCACTCGCGTATTGAAGG	$(AT)_6$	2 (Yes)	0.033	0.063	0.0494	DQ867090
DVV-ED5	F: TCCTGGTCCGTACGAGCTAC R: GGAATATATTTTATTACACCATGTGA	$(AT)_5(GTAT)_2$	2 (No)	0.066	0.063	1.0000	DQ867091
DVV-ED11	F: TGTTGGGGAAACATTGTTAGG R: ATCTAAGGGGAATGCGGAGT	$(GA)_9$	2 (Yes)	0.383	0.399	0.7504	DQ867092
DVV-ED13	F: GGGCAATGTTTTCATTAAGC R: ACTTTCTGTCCACGCCAGTT	$(TG)_7$	2 (No)	0.492	0.484	1.0000	DQ867093
DVV-ED14	F: CAGTGGATTTATGCAATCAGG R: CACCAATCACAAAAAGGATCG	$(TG)_8$	3 (Yes)	0.279	0.364	0.0338	DQ867094
DVV-ED18	F: AAAGATCAGACAGCGGGCTA R: CACAAATAAAACATTCTAGTTCGTAAA	$(TA)_6$	13 (No)	0.623	0.742	0.0150	DQ867095
DVV-ET1	F: ATGAAATGCCCGATGAAAAG R: TTCCAACATAGTTGTCATCATCC	$(TGA)_7$	4 (No)	0.689	0.642	0.3317	DQ867096
DVV-ET3	F: TCATCTTCTTCTCGTTCTTCTTCA R: TCTATCCTAGCCCATGCACA	$(TCT)_{10}$	5 (No)	0.787	0.734	0.6047	DQ867097
DVV-ET4	F: TAATTAACCCAGGCCACAGC R: AAACTGGGGGAAGCCTATGT	$(TTC)_7$	3 (No)	0.574	0.577	0.3619	DQ867098
DVV-ET6	F: AAGGCACCGTTGAAGAAAAA R: CCCTGAACAATCAATTGCAG	$(AAG)_5N_9(AAG)_{10}$	9 (No)	0.557	0.727	0.0049	DQ867099
DVV-ET10	F: TGGAGGCACATTGAAAAACA R: ATCGGTGGCAGGAAGATACA	$(AAG)_7$	2 (Yes)	0.339	0.488	0.0186	DQ867100
DVV-ET11	F: GGCTCGATAACACGGAAAGA R: TAAGTGGTCGATGGGTGGTT	$(ATA)_6$	2 (No)	0.033	0.032	1.0000	DQ867101
DVV-ET13	F: AACGTTTCTTTGGGCTAGTG R: TCTGCAAAATACAGGGTGGTC	$(TAT)_6$	4 (No)	0.607	0.570	0.6044	DQ867102
DVV-ET14	F: TAAGGACGTCCGTTCCAGAC R: TTGCCAGCATTATGGTCATC	$(ATG)_6$	5 (No)	0.459	0.444	0.0311	DQ867103
DVV-ET15	F: TTCAACAGAATGCTCTTGTCAAT R: ATTCATCGTCGATCCAGACC	$(ATG)_6$	2 (Yes)	0.213	0.345	0.0051	DQ867104
DVV-ET17	F: AAATTCCCCCAGGTCTATGG R: TGCCGTCTTCCTCTTGTACC	$(AGT)_6$	2 (No)	0.246	0.240	1.0000	DQ867105
DVV-ETet1	F: ATCGGTGGGTGTGAGGTATC R: TGTCCGACAAATCGTTCAAA	$(ATAC)_7$	2 (Yes)	0.459	0.497	0.6074	DQ867106

[a] Based on a survey of 10 controlled families.

[b] Hardy–Weinberg exact test, GENEPOP version 3.1b (Raymond and Rousset 1995).

females and *D. v. zeae* males (Krysan and Branson 1977). The TX population appeared morphologically to be *D. v. virgifera* but was collected in a putative hybrid zone in northwestern Texas (Krysan et al. 1980). For testing cross-species amplification, northern corn rootworm, *Diabrotica barberi*, and southern corn rootworm, *Diabrotica undecimpunctata howardi*, were collected in 2003 from eastern South Dakota and central Iowa, respectively. Adults of mixed sexes were stored at −80 °C until processed for DNA isolation. DNA was extracted from individuals using BioRad's Aqua Pure isolation kit (BioRad, Hercules, CA), according to the manufacturer's protocol.

SSRs

A database of 6397 EST sequences derived from the head of adult *D. v. virgifera* (S.T. Ractliffe and L. Liv, unpublished data) was screened for SSRs. The EST sequences were queried in a BLAST search with all possible combinations of core sequences of 5 dinucleotide repeats (total of 14 sequences), 5 trinucleotide repeats (total of 62 sequences), and 4 tetranucleotide repeats (total of 254 sequences). Of the 305 EST-SSRs identified, 133 were unique and had flanking regions suitable for primer design. Thirty-eight of these were selected for marker development based on having sequences containing di-, tri-, and tetranucleotides of at least 6, 6, and 5 repeat units, respectively. Of these, 17 produced discernable polymerase chain reaction (PCR) products and exhibited polymorphism in a sample of 60 individuals from an Iowa *D. v. virgifera* population and were characterized for inheritance, genetic diversity, and cross-species transferability (Table 1). The results were compared with those generated by 17 anonymous SSR markers isolated from genomic DNA using a biotin enrichment

and hybridization strategy (Kijas et al. 1994) as reported in Kim and Sappington (2005b). The loci were amplified in multiplexed PCR reactions, and individuals were genotyped using a Beckman-Coulter CEQ 8000 Genetic Analysis System using the methods described in Kim and Sappington (2005b).

Seven of the 17 EST-SSR loci exhibited at least 3 alleles in the survey of polymorphism in the IA population and showed no evidence of null alleles based on family analyses (described below). In addition, these 7 loci showed consistent readability and repeatability, appeared to be selectively neutral, and displayed no linkage between themselves. Together, these are the same criteria that were used to select the 7 genomic SSR loci, with the exception that a null allele was detected for *DVV-D*5 at very low frequency (0.025). Results of population genetics analyses based on the 7 EST-SSRs were compared with those generated for the same set of samples by the 7 genomic SSRs. Because the same criteria were used for screening both EST and genomic SSRs, any ascertainment bias is expected to be the same for the 2 classes of marker, making direct comparisons valid. The loci were amplified from 51 to 61 individuals per population in multiplexed PCR reactions, and individuals were genotyped as described above.

Cross-Taxa Amplification

Seventeen loci from each type of SSR were analyzed for transferability against a panel of 5 individuals each of Mexican, northern, and southern corn rootworms. Initially, PCR reactions were carried out using a single marker per reaction tube at an annealing temperature of 57 °C. If a clear product was not produced, PCR was attempted again using the touchdown method with a 1-degree reduction at each cycle from 61 °C to 52 °C, followed by 27 cycles at an annealing temperature of 52 °C.

Occurrence of Null Alleles

Ten parental crosses were set up in 2003 as single-pair matings between unrelated western corn rootworm adults from the main diapausing colony and the nondiapausing colony maintained at the Northern Grain Insects Research Laboratory in Brookings, SD. The nondiapausing colony had been in culture for about 190 generations and the diapausing colony for about 22 generations when the crosses were made (Kim et al. 2007). Previous analyses of SSR data from these colonies indicated that the genetic diversity in the nondiapausing line is roughly 25% less than that of contemporary wild populations but that there has been no detectable loss of diversity in the diapausing colony (Kim et al. 2007). Parents to be used in the crosses were sexed and segregated as pupae (George and Hintz 1966) to prevent premature mating. The male in each pair was caged with the female until the female began to lay eggs, after which the male was removed and stored at −80 °C. An aliquot of 100–400 eggs collected during the first 7 days after onset of oviposition by a given female was managed according to standard rearing procedures (Jackson 1986). Genomic DNA was extracted from the parents and ~50 of the F_1 offspring from each cross. To test for the presence of null alleles at each SSR locus, we 1) applied a parentage test, which is a simple comparison of offspring and parental genotypes and 2) checked for distortions to expected Mendelian segregation. Seventeen genomic and 17 EST-SSR loci were examined across 10 controlled families. With this number of families, one can expect to detect a null allele at a frequency of 0.072 with 95% confidence.

Population Genetic Analysis

Three estimates of genetic diversity were measured using the Microsatellite Toolkit (Park 2001), namely, allelic diversity which is the mean number of alleles per locus, observed heterozygosity (H_O), and unbiased estimates of expected heterozygosity (H_E) under Hardy–Weinberg assumptions for each population (Nei 1987). F statistics and pairwise F_{ST}'s (Weir and Cockerham 1984) were calculated using FSTAT v. 2.9.3 (Goudet 1995). Multiple comparisons were made after applying the sequential Bonferroni correction (Rice 1989). The exact probability approach (Guo and Thompson 1992) was used to test whether a population was in Hardy–Weinberg equilibrium (HWE), as implemented in GENEPOP (Raymond and Rousset 1995).

Genetic divergence based on allele frequency differences between populations was assessed with Nei's genetic distance (D_A; Nei et al. 1983) calculated by the program DISPAN (Ota 1993). The D_A matrix was used to construct phylogenetic trees by neighbor-joining clustering (Saitou and Nei 1987). The robustness of dendrogram topologies was tested by bootstrap resampling ($n = 1000$). A covariance matrix of allele frequencies across all loci was subjected to principal component (PC) analysis as applied in the program XLSTAT Addinsoft, NY. Factor scores were plotted in a scattergram along the 2 PC axes accounting for the most variation to facilitate visualization of geometric relationships among *D. v. virgifera* populations.

To compare the performance of the 2 types of SSRs in assigning individuals to their source population, we calculated the probability of each individual's belonging to a set of reference populations using both direct and simulation approaches implemented in the program GeneClass2 (Piry et al. 2004). The Bayesian statistical approach of Rannala and Mountain (1997) was employed for assignment/exclusion tests. For the exclusion test, we ran 10 000 Monte Carlo simulations of independent individuals for the population (Paetkau et al. 2004). When assigning individuals to a population, we followed the "leave one out" procedure (Efron 1983), wherein each individual was excluded from its own population.

The number of distinct populations (K) represented in the set of samples was estimated from genotype data for each of the 2 types of SSRs, using the program STRUCTURE 2.0 (Pritchard et al. 2000). This procedure calculates $\Pr(X|K)$ (the posterior probability of K) that the observed set of genotypes (X) would occur across a designated range of possible K's, in our case 1–6. The model simulated an admixture of individual ancestry and was run with correlated allele frequencies among populations (Pritchard et al. 2000). We used an initial burn-in of 100 000 iterations followed by 1 000 000 iterations.

Table 2. Cross-taxon amplification using primers for 17 SSRs derived from *Diabrotica virgifera virgifera* ESTs, based on survey of 5 individuals from each taxon

	Diabrotica virgifera zeae		*Diabrotica barberi*		*Diabrotica undecimpunctata howardi*	
Locus	PCR method[a]	Size range (number alleles)	PCR method[a]	Size range (number alleles)	PCR method[a]	Size range (number alleles)
DVV-ED1	57 °C	192–194 (2)	TD	—	TD	—
DVV-ED5	57 °C	244–246 (2)	TD	227–240 (2)	TD	—
DVV-ED11	TD	178–184 (3)	TD	—	TD	172–182 (3)
DVV-ED13	57 °C	224–226 (2)	57 °C	218–224 (3)	57 °C	234 (1)
DVV-ED14	57 °C	237–245 (3)	57 °C	267 (1)	TD	—
DVV-ED18	57 °C	231–235 (3)	57 °C	237 (1)	57 °C	231 (1)
DVV-ET1	57 °C	157–170 (3)	57 °C	167–170 (2)	TD	—
DVV-ET3	57 °C	206–263 (4)	TD	—	TD	—
DVV-ET4	57 °C	150–153 (2)	TD	—	TD	—
DVV-ET6	57 °C	162–222 (5)	57 °C, TD	NS	57 °C, TD	NS
DVV-ET10	57 °C	145–151 (2)	57 °C, TD	NS	57 °C, TD	NS
DVV-ET11	57 °C	225–237 (4)	57 °C, TD	NS	57 °C, TD	NS
DVV-ET13	57 °C	154–163 (2)	57 °C	164–176 (5)	TD	—
DVV-ET14	57 °C	155–170 (4)	TD	153–187 (6)	TD	—
DVV-ET15	57 °C	184–186 (2)	57 °C	173–188 (5)	TD	—
DVV-ET17	57 °C	196–202 (2)	TD	—	TD	—
DVV-ETet1	57 °C	227–239 (3)	57 °C	208–216 (2)	TD	—

—, no amplification; NS, nonspecific bands, where exact sizing of bands is difficult.

[a] PCR was first attempted with an annealing temperature of 57 °C. If unsuccessful, a touchdown (TD) protocol was attempted.

Five runs were performed for each value of K to assess consistency of $\Pr(X|K)$ estimates between runs.

Finally, Spearman rank-order correlation coefficients and Wilcoxon rank sum test were used to assess differences in allelic diversity, heterozygosity, pairwise F_{ST}'s, genetic distances, and population assignment derived from genomic versus EST-SSR data. These tests were conducted using Statistix 8 software (Analytical Software 2003).

Selective Neutrality Tests

Selective neutrality of each locus was examined based on Ewens (1972) sampling distribution of neutral alleles under the infinite-alleles model. The Ewens–Watterson homozygosity test (Watterson 1978) and the Ewens–Watterson–Slatkin exact test (Slatkin 1994, 1996) were carried out using allele frequency distribution, as implemented in Arlequin 3.1 software (Schneider et al. 2000). In these tests, the expected null distribution of the homozygosity statistic (F_{exp}) is generated by simulating random neutral samples and is compared with the homozygosity observed in the original sample (F_{obs}). If the null hypothesis of selective neutrality is rejected ($\alpha = 0.05$), an F_{obs}/F_{exp} ratio less than 1 implies balancing selection in favor of heterozygotes and a ratio greater than 1 implies directional selection in favor of advantageous alleles.

Results

Characterization of EST-SSRs from Western Corn Rootworm

Of the 305 SSRs identified in the BLAST search of the EST library, 110 were dinucleotide, 149 trinucleotide, and 46 tetranucleotide repeats. After eliminating loci that either were redundant or for which we were unable to design PCR primers, a total of 133 SSRs were available for testing, where 57 contained dinucleotide, 67 trinucleotide, and 9 tetranucleotide repeats. Seventeen of these EST-SSRs were genetically characterized using a sample of 60 western corn rootworm adults collected from a population in central Iowa in 2003 (Table 1). The number of alleles per locus ranged from 2 to 13, and expected heterozygosity values ranged from 0.063 to 0.742. After correction for multiple tests (adjusted significance [5%] threshold = 0.00294), none of the markers exhibited statistically significant deviation from HWE (Table 1). However, based on parentage analysis of 10 families, 6 loci (*DVV-ED1*, *DVV-ED11*, *DVV-ED14*, *DVV-ET10*, *DVV-ET15*, and *DVV-ETet1*) were found to harbor null alleles. Primer sets for all 17 EST-SSRs amplified DNA of the MCR, a subspecies of the western corn rootworm. Nine and 3 primer sets successfully cross-amplified DNA of the northern and southern corn rootworm, respectively (Table 2).

Comparison of Genetic Characteristics in Genomic versus EST-SSRs

Genetic diversity, inbreeding coefficient, interspecific transferability, and the occurrence of null alleles were examined for 17 loci from each type of SSR (Table 3). Based on a survey of the same 60 *D. v. virgifera* adults, EST-SSRs showed significantly lower levels of genetic diversity than genomic SSRs (allelic diversity: $U = 50$; $P = 0.0010$; heterozygosity: $U = 59$; $P = 0.0034$, Wilcoxon rank sum test), but the 2 types of SSRs showed similar values of inbreeding coefficient ($U = 119$; $P = 0.3891$,

Table 3. Genetic characteristics of 17 genomic and 17 EST-SSR loci developed from *Diabrotica virgifera virgifera*

		Genomic[a]	EST
Genetic character[b]	Allelic diversity	8.5	3.8
	H_O	0.585	0.402
	H_E	0.678	0.436
	F_{IS}	0.138	0.109
Interspecific transferability[c]	*Diabrotica virgifera zeae*	17	17
	Diabrotica barberi	6	9
	Diabrotica undecimpunctata howardi	2	3
Occurrence of null alleles[d]	Null	10 (58.8%)	6 (35.3%)
	No null	7 (41.2%)	11 (64.7%)

[a] Based on data from Kim and Sappington (2005b), except for the occurrence of null alleles which was examined in this study.

[b] Based on survey of 60 individual adults sampled in Ankeny, Iowa, 2003.

[c] Based on survey of 5 individuals from each taxon (see Table 2).

[d] Based on survey of 10 unrelated controlled cross families. Percentage occurrence for each type of marker in parentheses.

Wilcoxon rank sum test). Primer sets for the EST-SSRs exhibited a higher success rate in cross-amplification across 3 rootworm taxa compared with genomic SSRs. All loci from both types of SSR were amplified in the MCR, a subspecies of western corn rootworm. However, more of the EST-SSRs were successfully amplified in northern corn rootworm and southern corn rootworm than were the genomic SSRs (Table 3). Eleven (64.7%) of 17 EST-SSRs and 7 (41.2%) of 17 genomic SSRs displayed no evidence of null alleles in parentage analysis of 10 families.

Population Genetic Comparisons of Genomic versus EST-SSRs

Seven EST-derived and 7 genomic SSRs were further characterized in a survey of 346 individuals from 6 *Diabrotica* populations (Table 4). A total of 77 and 58 alleles were detected across the 7 genomic and EST-SSR loci, respectively. The number of alleles ranged from 4 to 20 in the genomic SSRs and from 3 to 16 in the EST-SSRs. Mean H_E ranged from 0.299 to 0.877 at the genomic SSR loci and from 0.464 to 0.767 at the EST-SSR loci (Table 4). For the 5 populations of *D. v. virgifera*, genomic and EST-SSRs showed very similar multilocus estimates of genetic diversity, F_{IS}, F_{ST}, and R_{ST}, and reflect genetic homogeneity of western corn rootworm populations across the US Corn Belt. There was no significant difference between EST-derived and genomic SSRs for interlocus variance of genetic differentiation (Bartlett's test of equal variances; F_{ST}: $P = 0.5379$; R_{ST}: $P = 0.2628$) or its median value (Wilcoxon rank sum test; F_{ST}: $P = 0.3159$; R_{ST}: $P = 0.1742$). Although the values for F_{ST} and R_{ST} for individual loci were substantially higher when *D. v. zeae* was included, multilocus estimates of genetic diversity, F_{IS}, F_{ST}, and R_{ST} were similar when calculated from either genomic or EST-SSR data.

At the population level, multilocus estimates of heterozygosity and inbreeding coefficient (F_{IS}) were very similar between the genomic and the EST-SSRs (Table 5). However, allelic diversity was significantly higher in the genomic SSRs than in the EST-SSRs ($P = 0.002$, Wilcoxon rank sum test). Among populations, there were no significant differences in measures of genetic diversity generated by EST-derived or genomic SSRs, including H_E (Kruskal–Wallis statistic = 0.2345, $P = 0.9987$ for EST-SSRs; K–W statistic = 0.6241, $P = 0.9869$ for genomic SSRs) and allelic diversity (K–W statistic = 0.8432, $P = 0.9742$

Table 4. Characteristics of individual SSR loci of both genomic and EST derivation for all 6 populations of *Diabrotica virgifera* and for the 5 *Diabrotica virgifera virgifera* populations excluding the *Diabrotica virgifera zeae* population (in parentheses)

Locus	Allele number	Heterozygosity	F_{IS}	F_{ST}	R_{ST}
Genomic					
DVV-D2	11 (11)	0.720 (0.748)	0.042^{NS} (0.049*)	0.089*** (0.006***)	0.149 (0.001)
DVV-D4	11 (9)	0.722 (0.704)	-0.006^{NS} (-0.011^{NS})	0.049*** (0.003**)	0.031 (0.007)
DVV-D5	4 (4)	0.299 (0.232)	0.156*** (-0.028^{NS})	0.113*** (-0.002*)	0.015 (-0.003)
DVV-D8	20 (20)	0.877 (0.882)	-0.000^{NS} (0.001^{NS})	0.049*** (0.016***)	0.032 (-0.002)
DVV-D9	7 (7)	0.575 (0.552)	0.0238^{NS} (0.011^{NS})	0.057*** (0.015***)	0.004 (0.003)
DVV-D11	18 (14)	0.826 (0.821)	-0.004^{NS} (-0.004^{NS})	0.025*** (0.007***)	0.005 (-0.002)
DVV-T2	6 (6)	0.529 (0.528)	0.007^{NS} (0.039^{NS})	0.032*** (0.011***)	0.008 (0.012)
Mean genomic	77 (71)	0.650 (0.638)	0.019^{NS} (0.010^{NS})	0.055*** (0.009***)	0.036 (0.002)
EST					
DVV-ET1	5 (5)	0.604 (0.661)	0.026^{NS} (0.015^{NS})	0.184*** (0.001^{NS})	0.447 (0.013)
DVV-ET3	10 (6)	0.716 (0.704)	-0.038^{NS} (-0.024^{NS})	0.041*** (0.010^{NS})	0.299 (0.016)
DVV-ET4	3 (3)	0.570 (0.587)	0.011^{NS} (0.044^{NS})	0.120*** (0.005^{NS})	0.219 (0.007)
DVV-ET6	13 (13)	0.767 (0.756)	0.053* (0.054*)	0.030*** (0.003**)	0.047 (-0.005)
DVV-ET13	5 (5)	0.571 (0.590)	0.024^{NS} (0.026^{NS})	0.152*** (0.013*)	0.116 (0.020)
DVV-ET14	6 (5)	0.464 (0.423)	0.091** (0.022^{NS})	0.110*** (0.003^{NS})	0.088 (-0.001)
DVV-ED18	16 (16)	0.728 (0.729)	0.101*** (0.089**)	0.006*** (-0.001^{NS})	0.028 (0.009)
Mean EST	58 (53)	0.631 (0.636)	0.037** (0.033*)	0.089*** (0.005***)	0.108 (0.006)

Probability that value is not different from zero; NS: not significant; *$P < 0.05$, **$P < 0.01$, and ***$P < 0.001$.

Table 5. Allelic diversity (AD), observed heterozygosity (H_O), expected heterozygosity (H_E) at HWE, and inbreeding coefficient (F_{IS}) values across 7 genomic and 7 EST-derived *Diabrotica virgifera* SSR loci at the population level

Population (sample size)	Genomic				EST			
	AD	H_E	H_O	F_{IS}	AD	H_E	H_O	F_{IS}
IA (61)	8.43	0.641	0.621	0.032^{NS}	6.14	0.638	0.614	0.039^{NS}
IL (60)	7.71	0.645	0.636	0.015^{NS}	5.43	0.635	0.588	0.075^{NS}
NY (61)	7.86	0.633	0.642	-0.015^{NS}	5.86	0.602	0.569	0.056^{NS}
TX (51)	8.57	0.700	0.678	-0.012^{NS}	5.71	0.660	0.672	-0.018^{NS}
DKS (58)	7.86	0.603	0.584	0.032^{NS}	5.86	0.642	0.636	0.010^{NS}
MCR (55)	7.00	0.706	0.665	0.059^{NS}	5.57	0.610	0.574	0.059^{NS}
Mean	7.91	0.650	0.637	0.019	5.76	0.631	0.609	0.037

Probability that value is not different from zero, NS: not significant.

for EST-SSRs; K–W statistic = 0.6553, $P = 0.9853$ for genomic SSRs).

Pairwise F_{ST} estimates among 6 *Diabrotica* populations were calculated based on the 2 types of SSR markers (Table 6). Pairwise F_{ST} estimates derived from the EST-derived SSRs were not significantly different from genomic SSRs ($U = 107$; $P = 0.8357$, Wilcoxon rank sum test). Apart from the MCR population, most population pairs showed little genetic differentiation. The MCR population exhibited considerable genetic differentiation from all western corn rootworms. Compared with the genomic SSRs, the EST-SSRs revealed higher estimates of pairwise F_{ST} between Mexican and western corn rootworm populations. Analysis of the EST-SSRs indicated significant differentiation of the TX population from all others, whereas analysis of genomic SSRs indicated no significant differentiation between TX and IL or DKS. In either case, the F_{ST} values were very low except in the comparison with the MCR population. There was a significant correlation in the rank order of pairwise F_{ST}'s and genetic distance estimates between the 2 types of SSRs but no significant difference among the central values of distance measures (Table 7). There was no significant correlation in rank order between the 2 types of SSRs for the measures of genetic diversity, that is, allelic diversity and H_E (Table 7).

Five EST and 3 genomic SSR markers showed evidence of nonneutrality in at least 1 population (Table 8). In all but 1 case (genomic marker *DVV-D8* in the exact test), the number of populations where neutrality was rejected for a particular marker was 2 or less. Two EST-SSRs and 1 genomic SSR marker showed evidence of nonneutrality in only 1 of the 2 neutrality tests. At loci with significant nonneutrality, the ratios of observed to expected homozygosity (F_{obs}/F_{exp}) averaged less than 1 across all populations, ranging from 0.641 to 0.891 in EST-SSRs and 0.738 to 0.788 in genomic SSRs. For the 7 EST and genomic SSR markers, the proportions of populations for which neutrality was rejected did not differ significantly (Wilcoxon rank sum test; $P = 0.4942$ for Ewens–Watterson homozygosity test and $P = 0.9796$ for Ewens–Watterson–Slatkin exact test).

Comparison of EST versus Genomic SSRs in Population Assignment

Population relationships visualized by PC analysis revealed that the EST-SSRs showed similar power in resolution for population clustering (Figure 1). The EST-SSRs accounted for 100% of the variation along axes 1 and 2 combined, whereas the first 2 axes for the genomic SSRs explained 90% of the total variation. For both types of SSR, the scatter diagram of factor scores clearly distinguished the *D. v. zeae* (MCR) population from the 5 *D. v. virgifera* populations along axis 1, which accounted for most of the variation. Relative positions of the 5 *D. v. virgifera* populations were different along axis 2. However, this axis accounted for only a small percentage of the total variation for both sets of markers, so the differences portrayed along it are biologically negligible.

Assignment tests for 346 individuals from 6 populations revealed that 46.2% (160 out of 346) of individuals were correctly assigned to their source population based on data

Table 6. Pairwise estimates of F_{ST} (Weir and Cockerham 1984) from 7 genomic (above diagonal) and 7 EST-derived *Diabrotica virgifera* SSRs (below diagonal) between populations

	IA	IL	NY	TX	DKS	MCR
IA	—	0.001^{NS}	-0.002^{NS}	0.013*	0.005^{NS}	0.137*
IL	0.006^{NS}	—	0.007^{NS}	0.020*	0.013^{NS}	0.142*
NY	0.002^{NS}	0.006^{NS}	—	0.018*	0.006^{NS}	0.141*
TX	0.012*	0.001^{NS}	0.017*	—	0.014*	0.102*
DKS	-0.004^{NS}	-0.000^{NS}	0.002^{NS}	0.006^{NS}	—	0.150*
MCR	0.231*	0.224*	0.264*	0.190*	0.230*	—

Probability of not being significantly different than zero after corrections for multiple comparisons; NS: not significant; $*P < 0.05$.

Table 7. Spearman rank correlation coefficients (r) and their probability values (P) between population diversity estimates from genomic and EST-derived *Diabrotica virgifera* SSRs, and Wilcoxon rank sum U probability value for testing difference between central values of each estimate

Measure	Spearman rank correlation		Wilcoxon rank sum probability
	r	P	
Allelic diversity	0.588	0.175	0.002
Heterozygosity	−0.029	1.000	0.394
Pairwise F_{ST}	0.740	0.002	0.836
Nei's D_A distance	0.892	0.000	0.152
Nei's standard distance	0.739	0.000	0.868

from the genomic SSRs and 38.4% (133 out of 346) based on the EST-SSRs. Because the low success rate of assignment is due mainly to the low genetic differentiation of western corn rootworm populations across the US Corn Belt (see Kim and Sappington 2005a), 3 representative populations, IA, TX, and MCR populations, exhibiting relatively high genetic differentiation, were analyzed alone in assignment/exclusion tests (Table 9). In all, 82.6% of individuals were correctly assigned to their source population using the genomic SSRs and 79.6% using the EST-SSRs. The difference in assignment success rates was not significant ($U = 11$; $P = 0.2963$ in Wilcoxon rank sum test).

Using the model-based Bayesian analysis for estimating the number of distinct populations (K), both types of SSR consistently indicated that the 6 locations sampled represented 3 distinct populations (posterior probability for $K = 3$: $P > 0.999$ for both types of SSR). Clustering of locations by STRUCTURE for both classes of marker showed clearly that MCR comprised 1 of the populations, but composition of the other 2 populations was ambiguous. However, the locations suggested by F_{ST} pairwise comparisons (Table 6) and population assignment tests (Table 9) as most likely comprising the 3 populations were the same for both types of SSR, such that population 1 consists of IA, IL, NY, and DKS, population 2 consists of TX, and population 3 consists of MCR.

Discussion

General Characteristics of EST-SSRs Relative to Genomic SSRs

General genetic characteristics of individual EST-SSRs from *D. v. virgifera* relative to genomic SSRs—that is, interspecific transferability, level of genetic diversity, and occurrence of null alleles—are comparable to previous findings from a wide range of species. Although it is possible that some of the cross-amplification products for both EST and genomic SSRs were from different regions of the genome, we assume that most cross-amplified loci are homologous given the similar size range of alleles observed. Our finding of enhanced cross-transferability of EST-SSRs from *D. v. virgifera* to other *Diabrotica* species relative to genomic SSRs is consistent with previous reports from many organisms, including shrimp (Perez et al. 2005; Wang et al. 2005) and fish (Coulibaly et al. 2005).

The mean heterozygosity for the 7 EST-SSRs across 6 populations was very similar to that for the 7 genomic SSRs, and 3 of the EST-SSRs showed a greater level of heterozygosity than the average heterozygosity of genomic SSRs. We found a lower level of allelic diversity in *Diabrotica virgifera* EST-SSRs compared with genomic SSRs, a pattern found in a parasitic nematode (Grillo et al. 2006) and commonly reported in plants (Cho et al. 2000; Scott et al. 2000; Eujayl et al. 2001; Leigh et al. 2003; Varshney et al.

Table 8. Result of selective neutrality test for the EST-derived and genomic SSR loci in 6 corn rootworm populations

SSR Type	Locus	Homozygosity statistic (F)			Neutrality rejected[a]	
		F_{obs}	F_{exp}	F_{obs}/F_{exp}	Ewens–Watterson test	Slatkin's exact test
Genomic	*DVV-D2*	0.287	0.347	0.826	0	0
	DVV-D4	0.284	0.361	0.788	0	2
	DVV-D5	0.705	0.667	1.057	0	0
	DVV-D8	0.131	0.172	0.762	2	3
	DVV-D9	0.430	0.531	0.810	0	0
	DVV-D11	0.181	0.245	0.738	1	1
	DVV-T2	0.476	0.605	0.786	0	0
EST	*DVV-ET1*	0.401	0.565	0.710	2	2
	DVV-ET3	0.290	0.452	0.641	1	0
	DVV-ET4	0.435	0.679	0.641	2	1
	DVV-ET6	0.240	0.336	0.715	0	2
	DVV-ET13	0.434	0.581	0.747	0	0
	DVV-ET14	0.541	0.515	1.051	0	0
	DVV-ED18	0.279	0.313	0.891	1	1

F_{obs}: sum of the squares of all allele frequencies; F_{exp}: expected F values under neutrality obtained by computer simulations (number of simulated samples = 16 000); Values of F_{obs}, F_{exp}, and F_{obs}/F_{exp} averaged across 6 populations.

[a] Number of populations out of 6 tested in which neutrality was rejected ($P < 0.05$).

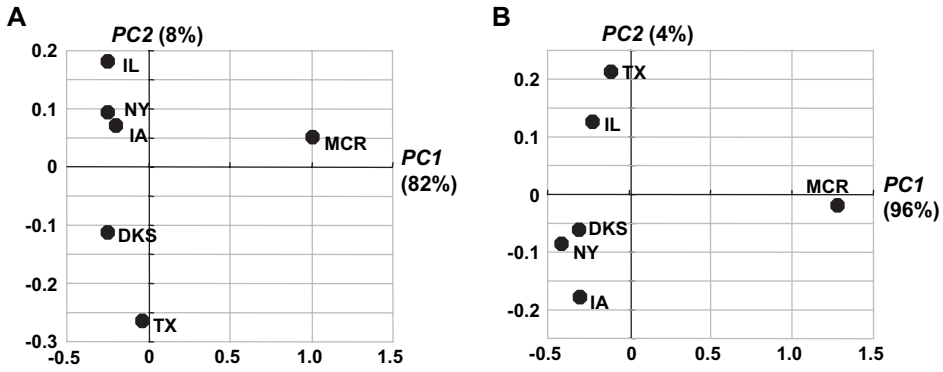

Figure 1. Scatter diagram of factor scores for 5 *Diabrotica virgifera virgifera* populations and 1 *Diabrotica virgifera zeae* population (MCR) derived from PC analysis of a covariance matrix of allele frequencies for 7 (**A**) genomic and (**B**) EST-SSRs. Percentage variation accounted for by each axis is indicated.

2005; Woodhead et al. 2005; Yatabe et al. 2007). Interestingly, equal or greater diversity has been reported for EST-SSRs in several species of fish (Liu, Tan, et al. 1999; Yue et al. 2004; Coulibaly et al. 2005; Cheng et al. 2007). The number of alleles and level of polymorphism of molecular markers influences the robustness of conclusions drawn from population genetic analyses (Kalinowski 2002), and there are a number of factors that can affect these parameters in EST-SSRs. Although functional constraints in transcribed regions of the genome appear in general to cause lower genetic diversity in nested SSRs, the level of polymorphism in EST-SSRs also is influenced by the type of unit repeats (i.e., di-, tri-, or tetranucleotide), the number of unit repeats, and the region of the gene in which they occur (Metzgar et al. 2000; Wren et al. 2000; Li et al. 2002; Coulibaly et al. 2005; Prasad et al. 2005; Varshney et al. 2005; Bouck and Vision 2007). The mean heterozygosity for the 10 trinucleotide EST-SSRs isolated in this study (H_E = 0.480) was substantially higher than that for the 6 dinucleotide plus 1 tetranucleotide loci (H_E = 0.373). This finding suggests that many of the EST-SSRs are in protein-coding regions because changes in trinucleotide repeat number will not cause frameshifts unlike changes in di- and tetranucleotide repeat numbers (Liu, Tan, et al. 1999; Wren et al. 2000; Li et al. 2004; Ellis and Burke 2007).

From parentage analyses of the same set of *D. v. virgifera* families, we detected fewer null alleles in EST-SSRs than in genomic SSRs. This was expected because functional constraints should reduce the frequency of null alleles in EST-SSRs (Ellis and Burke 2007). Null alleles arise from a mutation in the flanking region of the SSR locus, which prevents annealing by the associated PCR primer (Callen et al. 1993; Behura 2006). The presence of null alleles in SSRs biases estimation of population genetic parameters (Chapuis and Estoup 2007) and impacts molecular parentage analyses (Pemberton et al. 1995; Liewlaksaneeyanawin et al. 2002; Jones and Ardren 2003; Dakin and Avise 2004). The null allele problem is particularly conspicuous in insect SSRs. The frequency of null alleles has been reported as 33% in honeybees (Oldroyd et al. 1996), 50% in damselflies

Table 9. Percentage of *Diabrotica virgifera* individuals assigned to and excluded from (i.e., determined not to be a potential immigrant from) each reference population

		Potential source (reference) population					
		IA		TX		MCR	
Population (*n*)	Method[a]	Genomic	EST	Genomic	EST	Genomic	EST
IA (61)	Assignment	73.8 (45)	68.9 (45)	24.6 (15)	31.2 (19)	1.6 (1)	0.0 (0)
	Exclusion	3.3 (2)	4.9 (3)	0.0 (0)	4.9 (3)	95.1 (58)	95.1 (58)
TX (51)	Assignment	19.6 (10)	29.4 (15)	78.4 (40)	70.6 (36)	2.0 (1)	0.0 (0)
	Exclusion	21.6 (11)	5.9 (3)	0.0 (0)	3.9 (2)	82.4 (42)	90.2 (46)
MCR (55)	Assignment	1.8 (1)	0.0 (0)	1.8 (1)	0.0 (0)	96.4 (53)	100.0 (55)
	Exclusion	94.6 (52)	94.6 (52)	56.4 (31)	61.8 (34)	1.8 (1)	1.8 (1)

Assignment: the number of individuals assigned to the most likely population is in parentheses; Exclusion: the number of individuals excluded from the reference population at $\alpha = 0.01$ is in parentheses.

[a] Assignment test was carried out using the direct approach without probability computation, and the exclusion test was carried out using a simulation method (Cornuet et al. 1999). Both tests employed the Bayesian statistical approach of Rannala and Mountain (1997). The simulation method of Paetkau et al. (2004) was used in the exclusion test.

(Cooper et al. 1996), and 75% in white pine weevil (Liewlaksaneeyanawin et al. 2002). In this study with *D. v. virgifera*, we found that 59% of the genomic SSRs and 35% of the EST-SSRs harbored a null allele.

EST-SSRs for Characterizing Population Genetics

The results showed that population genetic parameters generated by the EST-SSRs are very similar to those obtained from the genomic SSRs for the same set of individuals from 6 populations. The low level of genetic differentiation among *D. v. virgifera* populations revealed by both sets of markers is consistent with previous findings of minimal genetic structuring among geographic populations in the US Corn Belt (Kim and Sappington 2005a; Miller et al. 2007). Genetic differentiation of the *D. v. zeae* population from the others corresponds to its subspecific status. The low but significant differentiation of the TX *D. v. virgifera* population may be related to its sympatry with *D. v. zeae* in the area of collection, where hybridization and limited introgression is suspected (Krysan et al. 1980; Giordano et al. 1997; Kim and Sappington 2005a). Estimates of pairwise F_{ST}'s and genetic distance measures generated by the EST-SSRs were similar to and significantly correlated with those produced by the genomic SSRs. Lack of significant rank correlations for allelic diversity and H_E can be ascribed to the general lack of significant differences observed in levels of these measures of genetic diversity among populations. Although allelic diversity was significantly greater for the genomic SSRs, H_E was similar between the 2 types of markers and thus quite sufficient for the EST-SSRs to be informative in population genetics applications.

Both types of SSR performed similarly in assigning individuals to their source population and in population clustering. In addition, EST-derived and genomic SSRs generated identical results in estimating the number of distinct populations (i.e., $K = 3$) represented in the total pool of individuals genotyped. Likewise, Walton et al. (2004) reported 100% correct assignment of individual mites to their original population based on EST-SSR data.

Tests for neutrality indicate that selection did not differentially affect performance of EST and genomic SSRs in characterizing *D. virgifera* populations. The grounds for this conclusion are 2-fold. First, there seems not to be strong selection acting on the 7 EST-SSRs used in this study. Instances of nonneutrality flagged by the tests were population specific, with all EST-SSR markers exhibiting neutrality in at least two-thirds of the 6 populations tested. Second, although deviation from selective neutrality was detected more often in EST-SSRs than in genomic SSRs based on the Ewens–Watterson test, the proportions of populations for which neutrality was rejected did not differ significantly. According to the Ewens–Watterson–Slatkin exact test, the number of instances where markers failed the neutrality test was the same for both EST and genomic markers. Even though EST-SSRs are potentially exposed to selection, Ellis and Burke (2007) point out that only a small percentage of genes in large-scale surveys (Tiffin and Hahn 2002; Clark et al. 2003) show evidence of positive selection. This is why allozymes usually behave as neutral markers, to the extent that they are universally recognized as legitimate markers for population genetics studies.

When *D. virgifera* EST-SSRs were used alone, the same or similar biological inferences were made for describing populations as when SSRs from noncoding, nontranscribed genomic regions were used. The observations made in this study are congruent with those drawn from the comparative study of genetic differentiation in fern populations (Woodhead et al. 2005), except that there was a significant rank correlation for allelic diversity in the latter. We conclude that the polymorphic SSRs derived from ESTs behave as effectively neutral markers in *D. virgifera* and are suitable for inferring population genetic structure in this insect. Together with the results of Woodhead et al. (2005) and Yatabe et al. (2007), the evidence suggests that EST-SSRs are appropriate markers for population genetics studies of both animals and plants. However, it is important to conduct selective neutrality tests on EST-SSRs before using them in population genetics analyses because even though most probably will not be under strong selection pressure, a small percentage may be (Ellis and Burke 2007; Yatabe et al. 2007). Increased speed and decreased expense of development, lower frequency of null alleles, and higher cross-species transferability are significant benefits of EST-SSRs that can be exploited in population genetics applications.

Acknowledgments

We thank Chad Nielson for expertly managing the crosses and rearing of families and Justin Gibson for excellent technical assistance in processing samples for DNA extraction. Mention of trade names or commercial products in this article is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the US Department of Agriculture.

References

Analytical Software. 2003. Statistix: user's manual. Version 8. Tallahassee (FL): Analytical Software.

Avise JC. 2006. The ontogeny of molecular ecology. Mol Ecol. 15:2687–2689.

Beaumont MA, Nichols RA. 1996. Evaluating loci for use in the genetic analysis of population structure. Proc R Soc Lond B. 263:1619–1626.

Behura SK. 2006. Molecular marker systems in insects: current trends and future avenues. Mol Ecol. 15:3087–3113.

Bouck A, Vision T. 2007. The molecular ecologist's guide to expressed sequence tags. Mol Ecol. 16:907–924.

Callen DF, Thompson AD, Shen Y, Phillips HA, Richards RI, Mulley JC, Sutherland GR. 1993. Incidence and origin of "null" alleles in the (AC)n microsatellite markers. Am J Hum Genet. 52:922–927.

Chapuis MP, Estoup A. 2007. Microsatellite null alleles and estimation of population differentiation. Mol Biol Evol. 24:621–631.

Chen C, Zhou P, Choi YA, Huang S, Gmitter FG Jr. 2006. Mining and characterizing microsatellites from citrus ESTs. Theor Appl Genet. 112:1248–1257.

Cheng L, Liao X, Yu X, Tong J. 2007. Development of EST-SSRs by an efficient FIASCO-based strategy: a case study in rare minnow (*Gobiocypris rarus*). Anim Biotechnol. 18:143–152.

Cho YG, Ishii T, Temnykh S, Chen X, Lipovich L, McCouch SR, Park WD, Ayres N, Cartinhour S. 2000. Diversity of microsatellites derived from genomic libraries and GenBank sequences in rice (*Oryza sativa* L.). Theor Appl Genet. 100:713–722.

Clark AG, Glanowski S, Nielsen R, Thomas PD, Kejariwal A, Todd MA, Tanenbaum DM, Civello D, Lu F, Murphy B, et al. 2003. Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios. Science. 302:1960–1963.

Cooper G, Miller PL, Holland PW. 1996. Molecular genetic analysis of sperm competition in the damselfly *Ischnura elegans* (Vander Linden). Proc R Soc Lond B. 263:1343–1349.

Cordeiro GM, Casu R, McIntyre CL, Manners JM, Henry RJ. 2001. Microsatellite markers from sugarcane (*Saccharum* spp.) ESTs cross transferable to erianthus and sorghum. Plant Sci. 160:1115–1123.

Cornuet JM, Piry S, Luikart G, Estoup A, Solignac M. 1999. New methods employing multilocus genotypes to select or exclude populations as origins of individuals. Genetics. 153:1989–2000.

Coulibaly I, Gharbi K, Danzmann RG, Yao J, Rexroad CE 3rd. 2005. Characterization and comparison of microsatellites derived from repeat-enriched libraries and expressed sequence tags. Anim Genet. 36:309–315.

Cruz F, Pérez M, Presa P. 2005. Distribution and abundance of microsatellites in the genome of bivalves. Gene. 346:241–247.

Dakin EE, Avise JC. 2004. Microsatellite null alleles in parentage analysis. Heredity. 93:504–509.

Efron B. 1983. Estimating the error rate of a prediction rule: improvement on cross-validation. J Am Stat Assoc. 78:316–331.

Ellis JR, Burke JM. 2007. EST-SSRs as a resource for population genetic analyses. Heredity. 99:125–132.

Ellis JR, Pashley CH, Burke JM, McCauley DE. 2006. High genetic diversity in a rare and endangered sunflower as compared to a common congener. Mol Ecol. 15:2345–2355.

Eujayl I, Sledge MK, Wang L, May GD, Chekhovskiy K, Zwonitzer JC, Mian MA. 2004. *Medicago truncatula* EST-SSRs reveal cross-species genetic markers for *Medicago* spp. Theor Appl Genet. 108:414–422.

Eujayl I, Sorrells M, Baum M, Wolters P, Powell W. 2001. Assessment of genotypic variation among cultivated durum wheat based on EST-SSRS and genomic SSRS. Euphytica. 119:39–43.

Ewens WJ. 1972. The sampling theory of selectively neutral alleles. Theor Popul Biol. 3:87–112.

Fagerberg AJ, Fulton RE, Black WC 4th. 2001. Microsatellite loci are not abundant in all arthropod genomes: analyses in the hard tick, *Ixodes scapularis* and the yellow fever mosquito, *Aedes aegypti*. Insect Mol Biol. 10:225–236.

Ford MJ. 2002. Applications of selective neutrality tests to molecular ecology. Mol Ecol. 11:1245–1262.

Fraser LG, Harvey CF, Crowhurst RN, De Silva HN. 2004. EST-derived microsatellites from *Actinidia* species and their potential for mapping. Theor Appl Genet. 108:1010–1016.

George BW, Hintz AM. 1966. Immature stages of the western corn rootworm. J Econ Entomol. 59:1139–1142.

Giordano R, Jackson JJ, Robertson HM. 1997. The role of *Wolbachia* bacteria in reproductive incompatibilities and hybrid zones of *Diabrotica* beetles and *Gryllus* crickets. Proc Natl Acad Sci USA. 94:11439–11444.

Goudet J. 1995. Fstat version 1.2: a computer program to calculate Fstatistics (Version 2.9.03). [cited 2008 Jan 3] J Hered. 86: 485–486. Available from: http://www2.unil.ch/popgen/softwares/fstat.htm.

Grillo V, Jackson F, Gilleard JS. 2006. Characterisation of *Teladorsagia circumcincta* microsatellites and their development as population genetic markers. Mol Biochem Parasitol. 148:181–189.

Grillo V, Jackson F, Cabaret J, Gilleard JS. 2007. Population genetic analysis of the ovine parasitic nematode *Teladorsagia circumcincta* and evidence for a cryptic species. Int J Parasitol. 37:435–447.

Guo SW, Thompson EA. 1992. Performing the exact test of Hardy-Weinberg proportion for multiple alleles. Biometrics. 48:361–372.

Jackson JJ. 1986. Rearing and handling of *Diabrotica virgifera* and *Diabrotica undecimpunctata howardi*. In: Krysan JL, Miller TA, editors. Methods for the study of pest *Diabrotica*. New York: Springer-Verlag. p. 25–47.

Johnson PCD, Webster LMI, Adam A, Buckland R, Dawson DA, Keller LF. 2006. Abundant variation in microsatellites of the parasitic nematode *Trichostrongylus tenuis* and linkage to a tandem repeat. Mol Biochem Parasitol. 148:210–218.

Jones AG, Ardren WR. 2003. Methods of parentage analysis in natural populations. Mol Ecol. 12:2511–2523.

Kalinowski ST. 2002. How many alleles per locus should be used to estimate genetic distances? Heredity. 88:62–65.

Kijas JM, Fowler JC, Garbett CA, Thomas MR. 1994. Enrichment of microsatellites from the citrus genome using biotinylated oligonucleotide sequences bound to streptavidin-coated magnetic particles. Biotechniques. 16:657–662.

Kim KS, French BW, Sumerford DV, Sappington TW. 2007. Genetic diversity in laboratory colonies of western corn rootworm (Coleoptera: Chrysomelidae), including a nondiapause colony. Environ Entomol. 36:637–645.

Kim KS, Sappington TW. 2005a. Genetic structuring of western corn rootworm (Coleoptera: Chrysomelidae) populations in the U.S. based on microsatellite loci analysis. Environ Entomol. 34:494–503.

Kim KS, Sappington TW. 2005b. Polymorphic microsatellite loci from the western corn rootworm (Insecta: Coleoptera: Chrysomelidae) and cross-amplification with other *Diabrotica* spp. Mol Ecol Notes. 5:115–117.

Kong Q, Xiang C, Yu Z, Zhang C, Liu F, Peng C, Peng X. 2007. Mining and charactering microsatellites in *Cucumis melo* expressed sequence tags from sequence database. Mol Ecol Notes. 7:281–283.

Krysan JL, Branson TF. 1977. Inheritance of diapause intensity in *Diabrotica virgifera*. J Hered. 68:415–417.

Krysan JL, Smith RF, Branson TF, Guss PL. 1980. A new subspecies of *Diabrotica virgifera* (Coleoptera: Chrysomelidae): description, distribution, and sexual compatibility. Ann Entomol Soc Am. 73:123–130.

Leigh F, Lea V, Law J, Wolters P, Powell W, Donini P. 2003. Assessment of EST- and genomic microsatellite markers for variety discrimination and genetic diversity studies in wheat. Euphytica. 133:359–366.

Li YC, Korol AB, Fahima T, Beiles A, Nevo E. 2002. Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. Mol Ecol. 11:2453–2465.

Li YC, Korol AB, Fahima T, Nevo E. 2004. Microsatellites within genes: structure, function, and evolution. Mol Biol Evol. 21:991–1007.

Liewlaksaneeyanawin C, Ritland CE, El-Kassaby YA. 2002. Inheritance of null alleles for microsatellites in the white pine weevil (*Pissodes strobi* [Peck] [Coleoptera: Curculionidae]). J Hered. 93:67–70.

Liewlaksaneeyanawin C, Ritland CE, El-Kassaby YA, Ritland K. 2004. Single-copy, species-transferable microsatellite markers developed from loblolly pine ESTs. Theor Appl Genet. 109:361–369.

Liu Z, Karsi A, Dunham RA. 1999. Development of polymorphic EST markers suitable for genetic linkage mapping of catfish. Mar Biotechnol. 1:437–447.

Liu Z, Tan G, Li P, Dunham RA. 1999. Transcribed dinucleotide microsatellites and their associated genes from channel catfish *Ictalurus punctatus*. Biochem Biophys Res Commun. 259:190–194.

Luikart G, England PR, Tallmon D, Jordan S, Taberlet P. 2003. The power and promise of population genomics: from genotyping to genome typing. Nat Rev Genet. 4:981–994.

Marcel TC, Varshney RK, Barbieri M, Jafary H, de Kock MJD, Graner A, Niks RE. 2007. A high-density consensus map of barley to compare the distribution of QTLs for partial resistance to *Puccinia hordei* and of defence gene homologues. Theor Appl Genet. 114:487–500.

Meglecz E, Petenian F, Danchin E, D'Acier AC, Rasplus JY, Faure E. 2004. High similarity between flanking regions of different microsatellites detected within each of two species of Lepidoptera: *Parnassius apollo* and *Euphydryas aurinia*. Mol Ecol. 13:1693–1700.

Metzgar D, Bytof J, Wills C. 2000. Selection against frameshift mutations limits microsatellite expansion in coding DNA. Genome Res. 10:72–80.

Miller N, Estoup A, Toepfer S, Bourguet D, Lapchin L, Derridj S, Kim KS, Reynaud P, Furlan L, Guillemaud T. 2005. Multiple transatlantic introductions of the western corn rootworm. Science. 310:992.

Miller NJ, Ciosi M, Sappington TW, Ratcliffe ST, Spencer JL, Guillemaud T. 2007. Genome scan of *Diabrotica virgifera virgifera* for genetic variation associated with crop rotation tolerance. J Appl Entomol. 131:378–385.

Navajas MJ, Thistlewood HMA, Lagnel J, Hughes C. 1998. Microsatellite sequences are under-represented in two mite genomes. Insect Mol Biol. 7:249–256.

Neff BD, Gross MR. 2001. Microsatellite evolution in vertebrates: inference from AC dinucleotide repeats. Evolution. 55:1717–1733.

Nei M. 1987. Molecular evolutionary genetics. New York: Columbia University Press.

Nei M, Tajima F, Tateno Y. 1983. Accuracy of estimated phylogenetic trees from molecular data. II. Gene frequency data. J Mol Evol. 19:153–170.

Neve G, Meglecz E. 2000. Microsatellite frequencies in different taxa. Trends Ecol Evol. 15:376–377.

Oldroyd BP, Smolenski AJ, Cornuet JM, Wongsiri S, Estoup A, Rinderer TE, Crozier RH. 1996. Levels of polyandry and intracolonial genetic relationships in *Apis dorsata* (Hymenoptera Apidae). Ann Entomol Soc Am. 89:271–283.

Ota T. 1993. DISPAN, genetic distance and phylogenetic analysis. Institute of Molecular Evolutionary Genetics, University Park, (PA): The Pennsylvania State University.

Paetkau D, Slade R, Burden M, Estoup A. 2004. Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. Mol Ecol. 13:55–65.

Park SDE. 2001. Trypanotolerance in West African cattle and the population genetic effects of selection [PhD thesis]. [Dublin (Ireland)]: University of Dublin.

Pashley CH, Ellis JR, McCauley DE, Burke JM. 2006. EST databases as a source for molecular markers: lessons from *Helianthus*. J Hered. 97:381–388.

Pemberton JM, Slate J, Bancroft DR, Barrett JA. 1995. Nonamplifying alleles at microsatellite loci: a caution for parentage and population studies. Mol Ecol. 4:249–252.

Perez F, Ortiz J, Zhinaula M, Gonzabay C, Calderon J, Volckaert FA. 2005. Development of EST-SSR markers by data mining in three species of shrimp: *Litopenaeus vannamei*, *Litopenaeus stylirostris*, and *Trachypenaeus birdy*. Mar Biotechnol. 7:554–569.

Piry S, Alapetite A, Cornuet JM, Paetkau D, Baudouin L, Estoup A. 2004. GeneClass2: a software for genetic assignment and first-generation migrant detection. J Hered. 95:536–539.

Prasad MD, Muthulakshmi M, Madhu M, Archak S, Mita K, Nagaraju J. 2005. Survey and analysis of microsatellites in the silkworm, *Bombyx mori*: frequency, distribution, mutations, marker potential and their conservation in heterologous species. Genetics. 169:197–214.

Primmer CR, Raudscpp T, Chowdhary BP, Møller AP, Ellegren H. 1997. Low frequency of microsatellites in the avain genome. Genome Res. 7:471–482.

Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. Genetics. 155:945–959.

Rannala B, Mountain JL. 1997. Detecting immigration by using multilocus genotypes. Proc Natl Acad Sci USA. 94:9197–9201.

Raymond M, Rousset F. 1995. GENEPOP population genetics software for exact tests and ecumenicism. (Version 1.2) J Hered. 86:248–249.

Rice WR. 1989. Analyzing tables of statistical tests. Evolution. 43:223–225.

Saitou N, Nei M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol. 4:406–425.

Sappington TW, Siegfried BD, Guillemaud T. 2006. Coordinated *Diabrotica* genetics research: accelerating progress on an urgent insect pest problem. Am Entomol. 52:90–97.

Schneider S, Kueffer JM, Roessli D, Excoffier L. 2000. ARLEQUIN: manual. Version 2.0. A software for population genetic analysis. Geneva (Switzerland): Genetics and Biometry Laboratory, University of Geneva.

Scott KD, Eggler P, Seaton G, Rossetto M, Ablett EM, Lee LS, Henry RJ. 2000. Analysis of SSRs derived from grape ESTs. Theor Appl Genet. 100:723–726.

Selkoe KA, Toonen RJ. 2006. Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. Ecol Lett. 9: 615–629.

Slate J, Hale MC, Birkhead TR. 2007. Simple sequence repeats in zebra finch (*Taeniopygia guttata*) expressed sequence tags: a new resource for evolutionary genetic studies of passerines. BMC Genomics. 8:52.

Slatkin M. 1994. An exact test for neutrality based on the Ewens sampling distribution. Genet Res. 64:71–74.

Slatkin M. 1996. A correction to the exact test based on the Ewens sampling distribution. Genet Res. 68:259–260.

Squirrell J, Hollingsworth PM, Woodhead M, Russell J, Lowe AJ, Gibby M, Powell W. 2003. How much effort is required to isolate nuclear microsatellites from plants? Mol Ecol. 12:1339–1348.

Stein N, Prasad M, Scholz U, Thiel T, Zhang H, Wolf M, Kota R, Varshney RK, Perovic D, Grosse I, Graner A. 2007. A 1,000-loci transcript map of the barley genome: new anchoring points for integrative grass genomics. Theor Appl Genet. 114:823–839.

Tiffin P, Hahn MW. 2002. Coding sequence divergence between two closely related plant species: *Arabidopsis thaliana* and *Brassica rapa* ssp. pekinensis. J Mol Evol. 54:746–753.

Varshney RK, Graner A, Sorrells ME. 2005. Genic microsatellite markers in plants: features and applications. Trends Biotechnol. 23:48–55.

Vasemagi A, Nilsson J, Primmer CR. 2005. Expressed sequence tag-linked microsatellites as a source of gene-associated polymorphisms for detecting signatures of divergent selection in Atlantic salmon (*Salmo salar* L.). Mol Biol Evol. 22:1067–1076.

Vigouroux Y, McMullen M, Hittinger CT, Houchins K, Schulz L, Kresovich S, Matsuoka Y, Doebley J. 2002. Identifying genes of agronomic importance in maize by screening microsatellites for evidence of selection during domestication. Proc Natl Acad Sci USA. 99:9650–9655.

von Stackelberg M, Rensing SA, Reski R. 2006. Identification of genic moss SSR markers and a comparative analysis of twenty-four algal and plant gene indices reveal species-specific rather than group-specific characteristics of microsatellites. BMC Plant Biol. 6:9.

Walton SF, Dougall A, Pizzutto S, Holt D, Taplin D, Arlian LG, Morgan M, Currie BJ, Kemp DJ. 2004. Genetic epidemiology of *Sarcoptes scabiei* (Acari: Sarcoptidae) in northern Australia. Int J Parasitol. 34: 839–849.

Wang H, Li F, Xiang J. 2005. Polymorphic EST-SSR markers and their mode of inheritance in *Fenneropenaeus chinensis*. Aquaculture. 249:107–114.

Watterson G. 1978. The homozygosity test of neutrality. Genetics. 88:405–417.

Weir BS, Cockerham CC. 1984. Estimating F-statistics for the analysis of population structure. Evolution. 38:1358–1370.

Woodhead M, Russell J, Squirrell J, Hollingsworth PM, Mackenzie K, Gibby M, Powell W. 2005. Comparative analysis of population genetic structure in *Athyrium distentifolium* (Pteridophyta) using AFLPs and SSRs from anonymous and transcribed gene regions. Mol Ecol. 14:1681–1695.

Wren JD, Forgacs E, Fondon JW 3rd, Pertsemlidis A, Cheng SY, Gallardo T, Williams RS, Shohet RV, Minna JD, Garner HR. 2000. Repeat polymorphisms within gene regions: phenotypic and evolutionary implications. Am J Hum Genet. 67:345–356.

Yatabe Y, Kane NC, Scotti-Saintagne C, Rieseberg LH. 2007. Rampant gene exchange across a strong reproductive barrier between the annual sunflowers, *Helianthus annuus* and *H. petiolaris*. Genetics. 175:1883–1893.

Yi G, Lee JM, Lee S, Choi D, Kim B-D. 2006. Exploitation of pepper EST-SSRs and an SSR-based linkage map. Theor Appl Genet. 114:113–130.

Yue GH, Ho MY, Orban L, Komen J. 2004. Microsatellites within genes and ESTs of common carp and their applicability in silver crucian carp. Aquaculture. 234:85–98.

Zane L, Bargelloni L, Patarnello T. 2002. Strategies for microsatellite isolation: a review. Mol Ecol. 11:1–16.

Zhang DX. 2004. Lepidopteran microsatellite DNA: redundant but promising. Trends Ecol Evol. 19:507–509.

Zhang DX, Hewitt GM. 2003. Nuclear DNA analyses in genetic studies of populations: practice, problems and prospects. Mol Ecol. 12: 563–584.

Received November 7, 2006
Accepted October 11, 2007

Corresponding Editor: Lacey Knowles